PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 1/13, 1/00, 1/107, C12P 21/02

A1

(11) International Publication Number: WO 95/03321

(43) International Publication Date: 2 February 1995 (02.02.95)

(21) International Application Number: PCT/US94/08127

(22) International Filing Date: 19 July 1994 (19.07.94)

(30) Priority Data:
08/094,819 20 July 1993 (20.07.93) US

(71) Applicant: BIONEBRASKA, INC. [US/US]; 3820 N.W. 46th Street, Lincoln, NB 68524 (US).

(72) Inventors: WAGNER, Fred, W.; Rural Route 1, Box 77B, Walton, NB 68461 (US). COOLIDGE, Thomas, R.; Beebe Hill Road, Falls Village, CT 06031 (US).

(74) Agent: BRUESS, Steven, C.; Merchant, Gould, Smith, Edell, Welter & Schmidt, 3100 Norwest Center, 90 South Seventh Street, Minneapolis, MN 55402 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD).

Published
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: METHOD FOR ENDOMODIFICATION OF PROTEINS

(57) Abstract

The invention provides a means for attaching a label, support or bioactive agent to the protein at a truncating site within the protein. More specifically the invention is directed to a method for the attachment of an amino acid, amino acid derivative, peptide or polypeptide, nucleophile to a protein core obtained from a protein of the core and a leaving group. In one embodiment, an endopeptidase enzyme is utilized to attach a labeled nucleophile to a protein core such as an antibody. In other embodiments, the invention is directed to a method for the attachment of a protein core to an immobilization support and to a method for the attachment of a bioactive agent to a protein core.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AT** | Austria | **GB** | United Kingdom | **MR** | Mauritania |
| **AU** | Australia | **GE** | Georgia | **MW** | Malawi |
| **BB** | Barbados | **GN** | Guinea | **NE** | Niger |
| **BE** | Belgium | **GR** | Greece | **NL** | Netherlands |
| **BF** | Burkina Faso | **HU** | Hungary | **NO** | Norway |
| **BG** | Bulgaria | **IE** | Ireland | **NZ** | New Zealand |
| **BJ** | Benin | **IT** | Italy | **PL** | Poland |
| **BR** | Brazil | **JP** | Japan | **PT** | Portugal |
| **BY** | Belarus | **KE** | Kenya | **RO** | Romania |
| **CA** | Canada | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CF** | Central African Republic | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CG** | Congo | | | **SE** | Sweden |
| **CH** | Switzerland | **KR** | Republic of Korea | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **KZ** | Kazakhstan | **SK** | Slovakia |
| **CM** | Cameroon | **LI** | Liechtenstein | **SN** | Senegal |
| **CN** | China | **LK** | Sri Lanka | **TD** | Chad |
| **CS** | Czechoslovakia | **LU** | Luxembourg | **TG** | Togo |
| **CZ** | Czech Republic | **LV** | Latvia | **TJ** | Tajikistan |
| **DE** | Germany | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **DK** | Denmark | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **US** | United States of America |
| **FI** | Finland | **ML** | Mali | **UZ** | Uzbekistan |
| **FR** | France | **MN** | Mongolia | **VN** | Viet Nam |
| **GA** | Gabon | | | | |

# METHOD FOR ENDOMODIFICATION OF PROTEINS

## Background of the Invention

It is well-known that the function of bioactive proteins can often be enhanced by their combination with other substances. When used to catalyze a reaction or to obtain separation, proteins can be immobilized to increase reaction efficiency and simplify the processing. When used as detecting agents, proteins can be labeled to facilitate measurement. When used to complex with or treat biological organisms, proteins can be combined with bioactive agents (hereinafter called "augmentation") to help achieve treatment efficacy.

Methods for immobilizing proteins are desirable because they localize reaction sites and improve economic recovery. Moreover, immobilized proteins are generally less susceptible to the loss of activity due to chemical attack and changes in temperature and pH than are free proteins.

Methods for labeling or augmenting proteins are desirable because they facilitate quantification, localization, specificity and reactivity of the protein. The resulting combinations, moreover, are the resulting combinations and/or powerful tools for clinical analysis and treatment.

Numerous techniques exist for protein immobilization on solid supports. Proteins can be physically adsorbed onto inert supports or can be covalently bound to the support through reaction with bifunctional linker arms. Microencapsulation, gel entrapment and complexation (with ion exchange resins) also can bind and immobilize.

Numerous techniques also exist for binding labels and bioactive agents to proteins. Most of these techniques call for reaction of the label or agent and a functional group of the protein, such as an amino group,

which occurs repeatedly throughout the protein. Although some repetitions of such groups are shielded from binding by the conformation of the protein, many others are exposed and available for binding with the labeling groups or bioactive agents. The result is a mixture of proteins having labels or bioactive agents attached at various non-specific sites.

One of the most important protein embodiments being investigated today is the antibody. The need to specifically attach immobilizing supports, labeling groups or bioactive agents to an antibody is widely recognized.

The known methods for immobilizing, labeling or augmenting any kind of protein fall far short of obtaining specific attachment at selected positions of the protein. Consequently, better and more specific methods for binding labels, supports or bioactive agents to proteins are needed.

Therefore, there is a need to develop methods for labeling, immobilizing or bioactively augmenting proteins at positional specific sites of the protein. There is also a need to immobilize proteins at positionally specific sites by covalently binding them to inert immobilization supports. Such needs also involve covalently binding a label or bioactive agents to a specific site of a protein.

## Summary of the Invention

The present invention is an economical biochemical method for endomodification of proteins by immobilizing, labeling and augmenting them through endopeptidase transpeptidation reactions. Generally, the method involves binding an immobilization support, label or bioactive agent (hereinafter called "auxiliary substance") to a protein core derived from a protein starting material composed of the core and a leaving group.

In particular, the method of the invention involves two primary reactions. The first (hereinafter called the "primary transpeptidation reaction"), catalytically transpeptidates an amino acid, a peptide, a polypeptide, or an amino acid derivative (hereinafter called "nucleophile"), to the protein core through the use of an endopeptidase enzyme. The second (hereinafter called the "primary binding reaction"), binds the side chain of the nucleophile to a specifically reactive group attached to the auxiliary substance.

The method of the invention can be practiced by alternative synthetic routes depending on whether the primary transpeptidation reaction or primary binding reaction is conducted first. These two routes are depicted in the scheme presented in the following section entitled Detailed Description of the Invention.

The first synthetic route can be employed with all sizes and solubilities of reactants. A nucleophile having a distinctive side chain is first transpeptidated to the selected cleavage site of the protein by the primary transpeptidation reaction. The transpeptidation involves contacting an endopeptidase enzyme, specific for an enzyme cleavage site, with a protein composed of at least one core linked by a cleavage site to a leaving unit, in the presence of a nucleophile. The endopeptidase enzyme cleaves the leaving unit from the core at the cleavage site and causes the core and the nucleophile to form an adduct of the core and nucleophile. The resulting protein-nucleophile adduct is then bound to auxiliary substance by the primary binding reaction which takes advantage of the distinctive character of the side chain. The adduct either is directly bound to the auxiliary substance or is indirectly bound through a bifunctional linker arm. In either case, the binding reaction occurs between the side chain of the nucleophile and a specifically reactive group on the auxiliary substance or linker arm.

The second synthetic route can be employed when the molarities of the reactants in the reaction medium are sufficient to permit relatively rapid enzymatic transpeptidation. A nucleophile is first bound to the auxiliary substance by the primary binding reaction to form an intermediate of auxiliary substance and nucleophile. The binding can be accomplished by direct reaction of the side chain of the nucleophile and the auxiliary substance or indirectly through a linker arm that has been prebound to the auxiliary substance. The intermediate is then transpeptidated to the protein core by the primary transpeptidation reaction described above. Transpeptidation with amino acid, peptide, polypeptide, or amino acid derivative nucleophiles can all be employed in this primary coupling reaction.

The protein starting material includes a core which may be a truncated version of its natural form. The core may be truncated through deletion of amino acids at either, or both, of its C-terminal and N-terminal ends depending on the product desired. The protein also includes a leaving unit linked to the core by an enzyme cleavage site recognized by the endopeptidase enzyme. The leaving unit may be one or more amino acid residues.

The protein starting material preferably is a biologically-active polypeptide. Included without limitation are enzymes, enzyme inhibitors, peptide hormones, DNA binding proteins, reading frame proteins, transcriptases, antibodies, $F_{ab}$ truncated antibodies, regulating proteins, peptides as small as two residues and various other functional proteins.

The protein starting material may as well be a fusion protein or a recombinant single copy polypeptide linked to a binding protein through an interconnecting peptide. The protein starting material may also be a recombinant multicopy polypeptide such as multiple copies of the single copy polypeptide linked together

with or without intraconnecting peptides. If an intraconnecting peptide is present, it preferably has at least one site that is selectively cleavable by the endopeptidase cleavage enzyme. The intraconnecting peptide may also serve as the leaving group from the C-terminal end of a single copy core polypeptide. The binding protein aids in the purification of the recombinant multicopy polypeptide.

The preferred proteins for use in the inventive method are monoclonal or polyclonal antibodies. The core constitutes the active fragments of such antibodies while the leaving group constitutes the constant region or some portion thereof. Preferred classes of antibodies include those that function to detect antigens in biological systems or contaminants in biological or inanimate systems, to carry bioactive agents to specific sites, to diagnose disease and organic disfunction, to separate antigens from other materials in biological or inanimate systems, and to remove antigens from biological or inanimate systems. Especially preferred embodiments are mammalian immunoglobulin proteins from the IgA, IgD, IgE, IgM, or IgG class of immunoproteins.

The amino acid, peptide, polypeptide and amino acid derivative nucleophiles used in the method of the invention are respectively (1) an alpha amino acid having a side chain with a reactive substituent, (2) a short sequence of alpha amino acid residues with one of the residues having a side chain with reactive substituent, (3) a moderate to long sequence of alpha amino acid residues with one of the residues having a side chain with a reactive substituent and (4) an amino acid residue to which is attached a synthetic, semi-synthetic or natural non-peptide organic group having a pendant group with a reactive substituent. The N-terminal amino acid or amino acid sequence of the nucleophile is constructed or selected to be a reactive

substrate for the endopeptidase. These kinds of nucleophile may have a simple, nonfunctional side chains in circumstances where they are also the auxiliary substance.

When the first synthetic route is employed, the side chain is chosen so that the nucleophile has a distinctive character relative to the amino acids of the protein. With this design, the nucleophile rather than the amino acids of the protein is selectively and preferentially reacted with a specifically reactive group of the linker arm or auxiliary substance. When the second synthetic route is employed, such a distinctive character can be used but is not necessary because the primary transpeptidation reaction provides the selectivity desired.

The linker arm used in the method of the invention is a flexible or semi-flexible chain which has as its termini (1) a specifically reactive group that is reactive with the side chain of the nucleophile and (2) an other functional group that reacts with a combining group of the auxiliary substance.

Immobilizing supports useful in the present invention are inorganic or organic materials which may be functionalized with a specifically reactive group for selective reaction with the side chain of the nucleophile, or with a combining group that reacts with the other functional group of the linker arm. The support is a porous or semiporous material that is biologically inert and insoluble in the medium used.

Bioactive agents include those that act to provide a desirable biochemical or therapeutic result. They may be functionalized with a specifically reactive group for reaction with the side chain of the nucleophile, or with a combining group that reacts with the other functional group of the linker arm. Included are chemotherapeutic agents, oxidizing or reducing agents, cytotoxic agents, anticancer agents, radioactive

agents, antibiotics, antimicotics, anti-infectives, heavy metal agents, antiviral agents, lysing agents, chelating groups and the like.

Labels useful in the present invention include fluorescent groups, phosphorescent groups, colorimetric groups, radioactive groups, luminescent groups, spectrometric groups, nuclear magnetic resonance groups, electron spin resonance groups and other groups with physicochemical properties that may be detected by measuring means. These labels may be functionalized with a specifically reactive group for reaction with the side chain of the nucleophile, or with a combining group that is reactive toward the other functional group of the linker arm. The nucleophile may also function as a label when it carries radioactive atoms.

The endopeptidase enzymes used according to the method of the invention include those of the serine or cysteine peptidase class. The endopeptidase enzymes trypsin and thrombin, of the serine peptidase class, and ficin and papin of the cysteine class, are especially desirable endopeptidase enzymes to serve as cleavage enzymes for the method of the invention.

The amino acid residue cleavage site for the endopeptidase enzyme may be recognized by the endopeptidase enzyme in solo or as a part of a multiple amino acid residue recognition sequence. In addition, according to the method of the invention, cleavage sites which are normally cleaved by an endopeptidase enzyme may be rendered less reactive or unrecognizable when adjacent to certain other amino acid residues. Use of this knowledge to render some cleavage sites less reactive is used advantageously to render substantial utility to endopeptidase enzymes which may otherwise be precluded from use in certain transpeptidation reactions. The ability to cause combination of the nucleophile with the core is a desirable characteristic of the endopeptidase enzyme.

The entire transpeptidation process may be done in a single step under very mild conditions. The number and sequence of steps of cleaving and reacting the starting material can vary depending on the starting material used. The conditions for the enzymatically catalyzed reaction between the protein and the nucleophile include control of pH, temperature, concentration and incubation time.

The present invention is also directed to methods which employ the labeled, immobilized or augmented protein.

The method for use of the labeled protein involves combining of the labeled protein and the material upon which it is to act, removing any excess labeled protein and measuring the amount of labeled protein that has interacted with the material. In particular, this method is useful for detection of antigens or enzymatic substrates/inhibitors by antibodies or enzymes, respectively.

The method for use of immobilized protein proceeds in a known manner as indicated by the character of the protein. The protein preferably is an enzyme, antibody, DNA binding protein or regulatory protein. The preferred uses will include enzymatically catalyzed reactions, antibody-antigen complexations, regulation of reactions and DNA or enzyme separations and/or purifications. One advantage of this method is the increased efficiency and ease of removal of the immobilized protein. Another advantage is the ability to increase the packing density of the immobilized protein when all molecules are aligned in the same direction and have exposed active sites.

The method for use of a bioactive agent bound to a protein also proceeds in a recognized manner as indicated by the bioactive agent and the nature of the protein. The action of the protein and bioactive agent cooperate to cause the effect desired. The protein may

act as a carrier to transport the agent across membranes or to cause its absorption into fluids, media or cells. It may also act as an absorption inhibitor to prevent transport of the agent across membranes or to prevent its absorption into fluids, media or cells. It may further act as a targeting vehicle to direct the agent to selective tissue sites or receptors. The advantage of this method is that by leaving the active sites free, the reaction efficiency and tissue selection are increased.

## Detailed Description of the Invention

Until the present invention, a general method for highly selective single site attachment of an auxiliary substance to a cleaved protein did not exist. The present methods solve this problem by providing precise control of the protein site to which the auxiliary substance is bound. This control causes the auxiliary substance to bind to a specific site on the cleaved protein.

The present invention provides a process for the selective modification of a protein by transpeptidation at cleavage sites specific for various cleavage enzymes. More specifically, the present invention is based upon the discovery that amino acids, peptides, polypeptides, and amino acid derivatives nucleophiles can be transpeptidated into proteins, especially biologically active proteins by a transpeptidation reaction under endopeptidase catalysis conditions (the primary transpeptidation reaction). The protein is composed of a core and a leaving unit such that during transpeptidation the nucleophile is substituted onto the core in place of the leaving unit. See J.S. Fruton in "Advances in Enzymology", A. Meister, ed. Vol. 53, 1982, John Wiley & Sons, New York, pp. 239-306 for a general review of proteinase catalyzed

synthesis of peptide bonds, the disclosure of which is incorporated herein by reference.

As described in the foregoing Summary of the Invention, the auxiliary substance is bound to the protein core through either of two synthetic routes. In the first, the nucleophile is separately transpeptidated to the protein core to form an adduct of the protein core and nucleophile. The adduct is then bound to the auxiliary substance directly, or is bound indirectly through a linker arm-auxiliary substance combination.

In the second synthetic route, the nucleophile and auxiliary substance are directly bound, or indirectly bound through a linker arm, to form an intermediate. The intermediate is then transpeptidated to the protein core.

**Primary Transpeptidation Reaction**

For purposes of this invention, "transpeptidation" is defined as that process whereby a terminal amino acid or a chain of amino acid residues (leaving unit), linked through an endopeptidase enzyme cleavage site at the C-terminal end of a protein core, is replaced by another amino acid, peptide, polypeptide, or amino acid derivative (nucleophile), in the presence of an endopeptidase cleavage enzyme. The method of the invention utilizes an endopeptidase enzyme, preferably of the serine or cysteine class, as the cleavage enzyme to catalyze the transpeptidation process.

The protein includes a core portion and a leaving unit. The core is any useful polypeptide sequence such as a native sequence, a modified native sequence, a non-native sequence preferably having biological activity, transacted forms thereof and similar versions. The leaving unit is linked to the core through an amide linkage and the amino acid residue or residues adjacent that linkage cause that linkage to be recognized as a cleavage site by the endopeptidase

cleavage enzyme. According to the method of the invention, the core linked to a leaving unit may be derived from any source including natural sources, semisynthetic-natural sources, manipulation of natural sources, chemical synthesis, or protein expression.

The protein is contacted with at least one endopeptidase cleavage enzyme specific for at least one cleavage site. The enzymatic cleavage of the protein at the linkage of the core portion and the leaving unit is conducted in the presence of a nucleophile.

The sequence and number of steps in the transpeptidation can be varied depending upon the desired modification of the protein, the amino acid sequence of the desired product peptide, and the starting material selected. The transpeptidation calls for the protein to be contacted with an endopeptidase cleavage enzyme, which has specific cleavage activity at the linkage between the core and the leaving unit.

The endopeptidase cleavage enzyme cleaves the leaving unit from the carboxy terminal of the core of the protein. Although it is not intended to be a limitation of the invention, it is believed that during this cleavage, the enzyme forms an acyl-enzyme intermediate with the core. In the presence of an appropriate nucleophile, under proper conditions, the enzyme causes the nucleophile to add to the cleaved core. Although it is not intended to be a limitation of the invention, it is believed that to accomplish this combination, the nucleophile displaces the cleavage enzyme from the acyl-enzyme intermediate and links to the protein core where the leaving unit was linked. The production of the modified protein is monitored by HPLC or other analytic procedure and the reaction is stopped by the addition of an acidic solution when the reaction has reached completion.

According to the method of the invention, preferably, the enzyme cleavage site recognition

sequence is not a sequence duplicated in the core or is not at an enzyme accessible sequence within the core.

The invention is further directed to modified enzyme cleavage sites which, when adjacent to certain amino acid residues, render the site unrecognizable or less reactive to cleavage. The discovery of the use of these unrecognizable or less reactive sites to prevent cleavage, renders a greater utility to various cleavage enzymes which would otherwise be precluded from use in certain transpeptidation reactions due to the detrimental effect of cleavage of proteins at sites within the desired core.

The leaving units to be cleaved from the core are specifically chosen to provide a suitable leaving unit for the specific endopeptidase cleavage enzyme. The nucleophiles are chosen to provide the amino acid or peptide chain to complete formation of the desired modified protein.

The two synthetic routes generally discussed above, are discussed in detail below. Both synthetic routes operate in two versions, which depend upon whether the auxiliary substance is directly or indirectly bound to the nucleophile.

**Route 1**

According to route 1, the first step is the primary transpeptidation reaction to form the adduct of protein core and nucleophile.

The choice of the particular nucleophile in the first step depends upon the identity of the amino acids of the protein and upon the distinctive character of the side chain of the nucleophile. The side chain of the transpeptidated nucleophile acts as the binding site for the specifically reactive group of the attached substance or the combination of linker arm-attached substance. It has a structure that either is non-duplicative of the amino acids of the protein or is more

highly reactive toward the specifically reactive group of the combination or auxiliary substance than are the amino acid side chains of the protein. It also is selected to avoid or minimize direct reaction with these side chains.

This selectivity imposed by the nucleophile side chain is accomplished by its reactive substituent. This substituent may be a sulfhydryl, olefinyl, amino, azidyl, hydrazinyl, epoxy, hydroxyl, activated hydroxy wherein the activator is a facile leaving group such as tosyl, mesyl and benzoyl, an acid group such as carboxyl, phosphoric or sulfonic, an activated ester such as a mixed anhydride, carbodiimido, iminyl amidinyl, imidazo, pivaloyl ester, neopentyl ester and the like, phosphoramidoyl, ferrocenyl, ferro complexes, boronyl and similar reactive functional groups.

In the second step of route 1, the primary binding reaction is accomplished by binding the nucleophile side chain either to the auxiliary substance or to its combination with the linker arm. In both of these variations, the specifically reactive group of the combination or auxiliary substance correlates with the reactive substituent of the side chain so that the side chain and auxiliary substance or combination readily react without substantially involving other groups of the protein.

To accomplish this selective reactivity of the primary binding reaction, the reactive substituent and specifically reactive group are correlated as pairs of groups. Several embodiments of this pair exhibit non-competitive binding which essentially will not involve other groups of the protein. These include, for example:

(1) a sulfhydryl and an organometallic group, preferably an organomercuric group or Alman reagent which are particularly useful with antibodies because antibodies do not naturally

contain free sulfhydryl groups, i.e. cysteine within or close to their active sites;

(2) an olefinyl group and a dienyl group, which form a Diels-Alder adduct;

(3) a phosphoramidoyl group and a metallophosphoramidoyl or metallophosphate group, which form co-ordinate complexes;

(4) an affinity complexing compound and its corresponding substrate, e.g. carbonic anhydrase and sulfanilamide or biotin and avidin, which form affinity complexes;

(5) a ferrocenyl group or ferro complex and a magnetic material rendered inert to the reaction medium, e.g. a teflon-coated iron wire coil, which form a magnetic transpeptidate;

(6) a chelating group and a chelated moiety such as ethylene diamine tetraacetate and a transition metal, which form a chelate;

(7) a polar olefinic or substituted olefinic group and the corresponding monomer, which polymerize by free-radical means under mild conditions, e.g. N-acryloyl lysine and acrylamide, and form a polymer;

(8) a pair of olefinic groups, which can be hydroborated and then treated with silver nitrate and weak base to form the reduced, transpeptidated olefin-olefin adduct; and

(9) a photoreactive arylketo group and a free radical stabilizing group having a radical-labile C-H bond; such as benzoylphenylalanine and a benzyl, allyl, or arylalkyl group, which can be photolyzed to form an adduct between the keto carbon and the C-H carbon of the free radical stabilizing group. Preferably such free radical stabilizing groups as polyamide, polycinnamide, polystyrene, or fluorene containing polymers (supports), porphyrin or

fluorescein, (label) and benzyl substituted bioactive agents are employed at significantly higher concentrations than the protein transpeptidated to the arylketo nucleophile so that photolytic addition of the protein to itself is highly disfavored (See J.C. Kauer, et al., J. Biol. Chem., 261, 10695 (1986)).

Other embodiments of this pair exhibit competitive binding relative to the functional groups of the protein but can be controlled to provide a substantially selective reaction of the side chain and attached substance or combination. These include, for example:

(1) an aromatic amino group and an epoxy, activated ester or aldehyde group, preferably an aromatic epoxy or aldehyde group, which can be reacted to form a nitrogen-carbon adduct, and under slightly acidic conditions to protonate the amine groups of the protein;

(2) an azidyl or hydrazinyl group and an aromatic amine, which can be reacted by irradiation with UV light to form a substituted amine, and under slightly acidic conditions to protonate the amine groups of the protein;

(3) an aromatic alcohol (e.g. phenolic group) or aromatic amine and an activated ester, which can be reacted to form an ester or amide respectively, and under slightly acidic conditions to protonate the amine groups of the protein; and

(4) a hydrazine and a reducing sugar, which form an osazone.

The conditions and procedures for performing the binding reactions of the side chain and specifically reactive group are known in the art. See for example "Reagents for Organic Synthesis" by Fieser & Fieser, John Wiley & Sons, New York, Vol. I-X, 1967-1975, the

disclosure of which is incorporated herein by reference. The conditions will generally be approximately ambient temperatures (0 to 38°C), and dilute to moderate concentrations of reactants. The procedures will generally involve stirred reactors, removal of side products and slow addition of reagents. A further condition is the maintenance of a minimal concentration of any reactant that can react with more than one group in the reaction mixture. For example, a minimum concentration of auxiliary substance or combination in the binding reaction is to be maintained so that the chance of undesirable side reactions of the specifically reactive group with the protein are minimized.

**Route 2**

According to route 2, the nucleophile and the auxiliary substance or its combination with linker arm are first bound by the primary binding reaction to form an intermediate. This step can be accomplished by employing (1) any of the reactive substituent and specifically reactive group pairs described above; (2) the combining group and other functional group pairs described below for the linker arm, or (3) by any of the known methods for forming an amide, ester, ether, imino, carbonate, urethane (carbamate), carbon-carbon, carbon-nitrogen, sulfur-carbon, sulfur-oxygen-carbon or carbon-oxygen bond. Methods to form these bonds and the particular groups formed thereby are known in the art. See, for example "Chemical Reagents for Protein Modification", CRC Press Inc., R.L. Lundblad & C.M. Noyes ed. 1984; "Basic Principles of Organic Chemistry", J.D. Roberts and M. Caserio, Benjamin Press, 1975, the disclosures of which are incorporated herein by reference.

Choice of the particular manner of binding the auxiliary substance or combination to the nucleophile does not depend upon the structure of the protein

according to this route. Any stable binding group that is appropriate for the chemical structures of nucleophile and auxiliary substance or linker arm combination will suffice because this binding reaction is not conducted in the presence of the protein.

The second step of route 2 transpeptidates the intermediate to the protein core through the primary transpeptidation reaction of the nucleophile portion of the intermediate and the protein (core and leaving unit). It is accomplished by endopeptidase catalysis under hydrolytic avoiding conditions. The selectivity of this reaction suits it as the one to be conducted in the presence of protein. The scheme of route 2 utilizes this feature to best advantage because it places the primary transpeptidation reaction last in the reaction sequence thereby eliminating the potential interference from the primary binding reaction.

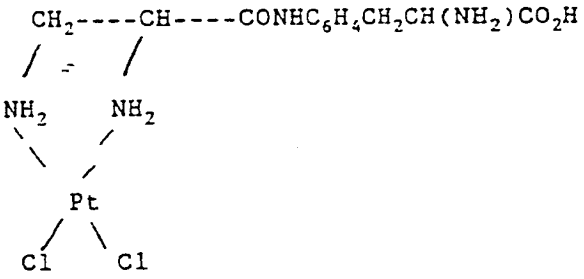
The second step of route 2 has some attendant parameters that primarily are directed to reaction efficiency. The reactants should have sufficient solubility in the reaction medium to enable relatively facile transpeptidation to take place. Generally, this solubility will be preferably about 0.05 to 2M for the reactants and 1 to 100 μM for the enzyme. When the solubilities of the reactants of the transpeptidation reaction are less than this, route 1 is preferentially employed.

Specific embodiments of labels, support materials and bioactive agents that can be transpeptidated to proteins by synthetic routes 1 and 2 are shown in Table 1.

Table 1

Amino Acids for Transpeptidation Labels, Bioactive Agents and Supports to Antibodies (Potential Applications are listed along with the chemical names of the Amino Acids)

| | | |
|---|---|---|
| $HOCH_2CH(NH_2)CO_2H$ | $H_2NCH_2CH_2SO_3H$ | $I(OH)C_6H_3CH_2CH(NH_3)CO_2H$ |
| Serine<br>$^3H, ^{14}C$<br>Radioactive Label | Taurine<br>$^3H, ^{14}C, ^{35}S$<br>Radioactive Label | Iodotyrosine<br>$^{129}I, ^{131}I$<br>Radioactive Label |
| $H_2NCH_2CH_2PO_3H$ | $FC_6H_4CH_2CH(NH_2)CO_2H$ | |
| Aminoethanephosphonic Acid<br>$^3H, ^{14}C, ^{32}P$<br>Radioactive Label<br>$^{31}P$ - NMR Label | Fluorophenylalanine<br>NMR Label | |




Metal Chelate
Platinum for Electron Microscopy and X-ray Labels
Cobalt for Radiation Therapy

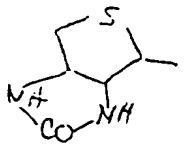
fluorescenyl $NHCSNH(CH_2)_4CH(NH_2)CO_2H$

Lysylfluorenscene
Fluorescent Label

Table 1 cont'd

$(NO_2)_2C_6H_3NH(CH_2)_4CH(NH)CO_2H$

Bisnitrophenyllysine
Fluorescent Label
Antigenic Label



$(CH_2)_4CONH(CH_2)_4CH(NH_2)CO_2H$

Biotin Group for Complex Attachment to Avidin-Resin or as Label for Enzymatio Detection

$H_2C=HCH_2NH$ + $CH_3CH=CHCH=CHCO_2$-(attached substance)

Diels-Alder
For Attachment of Resin
or other attached substance

$H_2C=CHCONH(CH_2)_4CH(NH_2)CO_2H$ + $H_2C=CHCONH$-(H or attached substance)

Free Radical Polymerization

$F_eC_5H_4$-$C_5H_3NH$- + wire

Magnetic Process

**Linker Arm**

There are several reasons why the linker arm version of routes 1 and 2 would be selected according to the invention. First, the environment of the protein may prevent approach of a large, bulky auxiliary substance. Second the auxiliary substance may not contain functional groups that are specifically reactive with the side chain of the nucleophile. Third, the linker arm increases the distance between the auxiliary substance and the protein which can help maintain the activity of the protein. Fourth, the protein can more freely adopt a spatial conformation that is appropriate for its reactivity. Fifth, it will lessen or minimize alteration of protein confirmation caused by auxiliary substance proximity.

The structure of the linker arm includes functional groups at the ends of a flexible to semi-flexible chain. One of the functional groups is the specifically reactive group mentioned above that reacts with the side chain of the nucleophile. The other functional group of the linker arm is chosen to readily react with available combining groups on the auxiliary substance. Of course, this pair of groups of the linker arm is selected so that one does not substantially interfere with the other when they are used in the binding and combining reactions.

In both synthetic routes, the step to combine linker arm and attached substance is accomplished before the binding reaction with the nucleophile. Since protein and nucleophile are not present during this step, the kinds of reactions available are numerous. If the combining group of the auxiliary substance is an aldehyde group, the other functional group may be an amine (Schiff base product), an activated acid such as an iminocarboxy, carboxyalkoxy or acid halide (amide product) or epoxy (substituted amine product). If the combining group of the auxiliary substance is a hydroxy

group, the other functional group may be an activated acid or ester (ester product) or activated alkyl such as a halo alkyl, alkyl tosyl or alkyl mesyl (ester product). If the combining group of the auxiliary substance is an acid group, the other functional group may be an amine (amide product) or activated hydroxyl (ester product). If the combining group of the auxiliary substance is a chelating agent, the other functional group may be a bound metal group. Other pairs of reactants include water soluble carbodiimide and amino; N-acyl succinimide and amino; and olefin and diene as well as those described above under part (3) of the primary binding reaction for route 2. Of course, the reverse order of reaction is also possible.

The backbone of the linker arm may be any that provides a flexible or semi-flexible chain. Included are polymers and oligomers of amides (peptides), olefins, esters, carbonates, urethanes, ethers, epoxides, and the like. Also included are alkylene and hydrocarbon chains. The length of the backbone may be from about two to about 100 atoms or monomeric units, preferably about four to about 20 atoms or monomeric units in length. Examples of the backbone include hexylenyl, decylenyl, poly(4-aminobutyric acid) poly(glycyl), poly(glycyl-alanyl), poly(4-hydroxybutyric acid), polylactones, poly(bisphenol-A-digylcidyl ether) and polyacrylamide.

**Proteins**

Most types of proteins (including peptides as small as two residues) can be transpeptidated according to the present invention. Examples of suitable proteins include enzymes, enzyme inhibitors, hormones including peptide hormones, antibodies, $F_{ab}$ truncated antibodies, functional proteins, transcriptases, reading frame proteins, DNA binding proteins, and other biologically active polypeptides.

With respect to labeling, monoclonal or polyclonal antibodies, DNA binding proteins, enzymes, and reading frame proteins are preferred as proteins. They are generally useful in the diagnosis of diseases, disorders, or hereditary dysfunctions. The antibodies are also generally useful in separation techniques and for detection of antigenic material. This includes mammalian immunoglobulin proteins from the IgA, IgD, IgE, IgM, or IgG class of immunoproteins.

With respect to immobilization, mono or polyclonal antibodies and enzymes are preferred as proteins.

With respect to bioactive agents, mono or polyclonal antibodies, peptide hormones, histocompatibility proteins, polypeptide inhibitors, peptide toxins, structural proteins (e.g. collagen), globular proteins and fibrous proteins are preferred as proteins.

**Nucleophile**

A wide variety of nucleophiles can be used in the present invention. Generally these include alpha amino acids or amino acid sequences (peptides, polypeptides) with neutral, basic, or acidic side chains wherein the side chains may contain the reactive substituents mentioned above. The choice of nucleophile for labeling antibodies, binding proteins to immobilizing support, or binding bioactive agents to proteins is generally coordinated with the enzyme and the protein chosen. For example, trypsin will generally transpeptidate nonacidic amino acids onto Arg or Lys residues of a protein. Certain other enzymes, which utilize cysteine and/or serine at their enzymatic sites and are derived from plant and microbial sources, will transpeptidate amino acids with acidic side chains to proteins.

Furthermore, in order to perform the binding reaction, the amino acid nucleophile will exhibit a distinctive reaction character. This will allow its selective binding to the auxiliary substance or its combination with linker arm. The distinctive character results from the reactive substituent of the side chain of the nucleophile as explained above. This reactive substituent may be a sulfhydryl, hydroxy, activated hydroxyl, phosphoramidoyl, hydrazinyl, amino, azidyl, epoxy, acid, boronyl, activated esters, ferrocentyl, ferro complex, or olefinyl group, mixtures thereof and other functionally reactive groups.

Embodiments of these amino acid nucleophiles include aliphatic amino acids such as monoamino monocarboxylic acids, e.g., glycine, alanine, valine, norvaline, leucine, isoleucine, and norleucine (useful as radioactive label); hydroxy amino acids such as serine, threonine, and homoserine; sulfur-containing amino acids such as methionine, cystine, cysteine, and taurine (for linker arm or auxiliary substance binding); diamino monocarboxylic acids such as orthinine, lysine, and arginine (for linker arm or auxiliary substance binding); and monoamino dicarboxylic acids such as aspartic acid and glutamic acid (for linker arm or auxiliary substance binding). Also aromatic amino acids, such as phenylalanine and tyrosine; heterocyclic amino acids, such as histidine and tryptophan, and olefinic amino acids such as 2-amino-2-vinyl acetic acid (for linker arm or auxiliary substance binding) are included within the group of amino acids of the present invention.

Additional amino acids are those with the C-terminal end protected. This includes, for example, amides, anilides, hydrazides, esters, and the like.

Embodiments of peptide nucleophiles include various sequences of the foregoing amino acid residues as well as any about two to about 50 unit sequence of

the twenty naturally occurring amino sequence. Examples include Gly Arg Mea or Gly-Biotin where Mea is maleic acid.

Preferred classes of nucleophile include aliphatic amino acids, hydroxy amino acids, their activated derivatives, phosphoramidoyl amino acids, sulfur-containing amino acids, diamino monocarboxylic acids, activated ester amino acids, aromatic amino acids, and heterocyclic amino acids. Especially preferred embodiments include serine, alanin, phenylalanine, taurine, lysine, arginine, 2-aminopenta-4-enoic acid and cysteine. Also included are carboxyl protected amino acids, such as amides and esters.

**Labels**

Labels for the proteins according to the present invention include labeled, or tagged amino acids having a variety of substituents or atoms that possess properties suitable for detection by conventional techniques. Such properties include photoaffinity, magnetism, radioactivity, fluorescence, enzymatic activity, electron dense (x-ray), nuclear magnetic resonance, electron spin resonance, antigenicity, and phosphorescence. For example, amino acids can be labeled with either $^{14}C$ or $^{3}H$ atoms. Further, the amino acids may be tagged by known fluorescent dyes, porphyrins, colorimetric dyes, reactive groups and antigens or enzymatic substrates that permit spectroscopic, photographic or radiometric detection. See E.T. Koh, et al., Biotechniques, 7, 596 et seq. (1988); S. Borman, "Bioconjugate Chemistry Attracts Growing Interest" in the May 8, 1989 issue of "Chemical and Engineering News" at p. 25 et seq., the disclosures of which are incorporated herein by reference.

Immobilizing Supports

Immobilizing supports useful in the present invention are inorganic or organic materials functionalized so that a reaction can occur between the nucleophile or linker arm and the support. When the former reaction is employed, the support will be functionalized with a specifically reactive group mentioned above. When the latter reaction is employed, the support will be functionalized with the combining group for the other functional group mentioned above. In this case also, the reactive substituent may be chelating ferromagnetic groups. The immobilizing support then has the appropriate character to produce binding. With a ferromagnetic group, the support may be magnetic wire that is rendered inert to the reaction medium, e.g. with teflon. Passing a current through the wire will establish the magnetism needed to cause binding. Alternatively, a magnet external to the system (i.e. outside the chromatographic medium) can be used to cause binding to the support. With the chelating group, the support may be an immobilized metal or other chelate.

The support may be a porous or semiporous solid. Preferably, it is biologically inert and insoluble. Materials that may be used as supports include fibers, sheets, microspheres, particles, beads, membranes, and the like.

The surface of the immobilizing support of the present invention is preferably porous. The use of substances having a porous surface, such as substantially spherical polymeric beads or microspheres of agarose allows large surface areas for the attachment of protein at high density. A surface is considered porous where the size of the majority of the pores in the material is sufficiently large so as to allow the migration of the protein into the interior of the spheres. The size and shape of the support may be

varied widely, depending on the particular protein and its intended use.

The immobilizing supports include a wide variety of substances. The choice of support, however, depends upon the choice of the nucleophilic and/or linker arm as well as on the intended use of the immobilized protein. The transpeptidation reactions, nucleophile, specifically reactive group and reactive group all are compatible as described above. In particular, the support is chosen such that the nucleophile will readily transpeptidate to the support or support-linker arm combination in preference to any other reactive sites on the protein. For example, cysteine may be used as the amino acid nucleophile to transpeptidate with a protein with no sulfhydryl groups, e.g. an antibody. A support or support-linker arm specifically reactive group is chosen that would react with the sulfhydryl moiety, for example, an organometallic group such as an organo mercury compound. Alternatively, 2-amino-hex-4-enoic acid may be the amino acid nucleophile, and a specifically reactive group for the support may be one that would specifically react with the unsaturated side chain, as for example through a Diels Alder reaction. Another alternative is the choice of a photoaffinity label such as N-hydroxy succinimidyl-4-azidosalicylic acid side chain, and an arylamine as the specifically reactive group on the attached substance. This salicylic side chain is to be transpeptidated to the epsilon amino group of a lysine before the photo addition so that it will not be reactive with the amino groups of the protein. Photoreaction under, for example, u.v. light, will accomplish the desired photo binding reaction. Moreover, if a linker arm is used, available groups on the support act as the reactive group. The other functional group of the linker arm is appropriately chosen to bind with the reactive group.

**Bioactive Agents**

Included within the invention is a method for attachment of a bioactive agent to a protein at a site remote from the active site. These bioactive agents can be carried or transported by the protein to a site where they can perform a desired reaction.

The bioactive (biologically active) agent includes physiologically or pharmacologically active substances that act locally or systemically in the body. Examples of biologically active agents include peptide drugs, protein drugs, desensitizing agents, antigens, vaccines, anti-infectives, antibiotics, antimicrobials, antiallergenics, steroidal anti-inflammatory agents, decongestants, miotics, anticholinergics, sympathomimetics, sedatives, hypnotics, psychic energizers, tranquilizers, androgenic steroids, estrogens, progestational agents, humoral agents, prostaglandins, analgesics, antispasmodics, antimalarials, antihistamines, cardioactive agents, nonsteroidal anti-inflammatory agents, antiparkinsonian agents, antihypertensive agents, β-adrenergic blocking agents, nutritional agents, metal compounds, anti-cancer compounds such as fluorinated nucleotides, nucleotide analogs, cytosine arabinocide, 5-fluorouracil, ricin-A, tetanus toxin, cyclic therapeutic peptides such as anamycin, erythromycin, cyclosporin, AZT, and alkaloids. Also, various forms of the biologically active agents may be used. Forms such as uncharged molecules, molecular complexes, salts, ethers, esters, and amides are included.

The bioactive agents are functionalized to carry specifically reactive groups for transpeptidation to the nucleophile directly. Alternatively, appropriate available combining groups on the bioactive agent can be reacted with the other functional group on a linker arm.

Preferably, this functionalization will be accomplished with a group already present within the agent.

**Cleavage Enzymes**

The cleavage enzymes, according to the method of the invention, include the class of endopeptidases. The endopeptidases suitable for use in the present invention include the serine and cysteine peptidases. Although it is not intended to be a limitation of the invention, the mechanism of action of serine and cysteine endopeptidases is believed to involve the formation of an acyl-enzyme intermediate with the core after cleaving the leaving unit. Under appropriate reaction conditions, it is believed that the nucleophile acts as a nucleophile and displaces the endopeptidase cleavage enzyme from the acyl-enzyme intermediate.

Serine peptidases are in a group of animal and bacteria endopeptidases which have a catalytically active serine residue in their active center. Representative examples of endopeptidases of the serine peptidase classification include trypsin and thrombin.

The endopeptidase trypsin is found in the pancreas of all vertebrates. It is released via the pancreatic duct into the duodenum as a trypsinogen. Conversion of trypsinogen into trypsin is initiated in the small intestine by enterokinase. Natural or synthetic forms of trypsin are suitable for the method of the invention.

Trypsin is known for its pronounced cleavage site specificity, catalyzing hydrolysis of only Lys-X and Arg-X (X is another amino acid residue) bonds. Trypsin's affinity for cleavage at the Lys-X bond is significantly diminished when immediately adjacent to an amino acid containing a carboxylic acid side chain, specifically including the amino acids glutamic acid and aspartic acid. A discovery of the present invention utilizes the knowledge of decreased cleavage activity at

the Lys-X cleavage sites which are adjacent to an amino acid containing a carboxylic acid side chain. This discovery has rendered the endopeptidase trypsin of great utility in the formation of modified proteins, according to the method of the invention. Natural or synthetic forms of thrombin are suitable for the method of the invention.

The glycoprotein endopeptidase thrombin, also of the serine peptidase classification, is responsible for the conversion of fibrinogen to fibrin. It is naturally produced during blood coagulation by the action of factor $X_a$ upon prothrombin. This endopeptidase has considerable sequence homology with trypsin and contains the catalytically important residues His, Asp, and Ser in the B chain. Thrombin has a cleavage specificity for the Arg-X cleavage site in specific peptide sequences known as recognition sites.

Thrombin is known for its cleavage site specificity at the carboxyl side of the Arg- residue. The known recognition sequence for the cleavage site Arg-X is Gly-Pro-Arg. A discovery of the present invention is that thrombin also cleaves at the carboxyl side Arg- residue within the recognition sequence of Gly-Ala-Arg. This discovery enhances the use of thrombin for transpeptidation by the method of this invention as well as other synthetic reactions where knowledge of the Gly-Ala-Arg recognition sequence will be of benefit.

**Transpeptidation Using the Endopeptidase Trypsin**

The transpeptidation process according to the method of the invention may be accomplished using starting protein derived from single or multicopy constructs, or single or multicopy fusion protein constructs.

## 1. Trypsin Transpeptidation Details

### Conditions for Primary Transpeptidation Reaction

The conditions for the primary transpeptidation reaction efficiently favor transpeptidation over peptide cleavage. As can be seen from the following discussion of the application of the primary transpeptidation reaction to antibody protein, these conditions generally involve control of Ph, temperature, reactant concentrations, enzyme concentration and incubation time.

The conditions for transpeptidation are within range of about pH 2 to 11. The selection of condensation over transpeptidation is made kinetically in that transpeptidation is completed quickly. In particular, transpeptidation occurs between 5 seconds and 1.4 hours.

The reaction temperature is the functional range of the enzyme, preferably up to about 40°C.

The concentrations of the reactants and enzyme are adjusted to provide optimum results. Generally, the highest possible concentrations of enzyme, nucleophile and protein are used that coincide with an appreciable primary transpeptidation reaction rate. Preferably, the protein is present at a concentration of from about 1 μM to about 1M, especially up to about 1 μM when the protein is an antibody. The nucleophile or intermediate incorporating the nucleophile is preferably present at a concentration of at least 0.05 molar and especially a concentration of from about 0.1 to 2 molar. The enzyme is preferably present at a concentration of about 1 to 100 μM, preferably about 1 to 100 μM.

The transpeptidation can also be facilitated by conducting the reaction in substantially all organic solvent and a minimum of water, e.g., 5-20% water. Useful organic solvents include DMF, DMSO, dimethyl acetamide, 1,4-butanediol and the like.

The incubation time (reaction time) of the protein and the nucleophile is from about 0.2 to 10 hours, preferably from 1.0 to 8 hours for condensation while for transpeptidation or transesterification, it is from about 30 seconds to about 1.5 hours.

The following description is based upon a particular core starting protein, however, it is understood that the method of the invention is suitable for transpeptidation of proteins, regardless of the source.

**Trypsin Transpeptidation of a Single Copy Recombinant Polypeptide**

The following description is based upon a particular recombinantly-derived core starting polypeptide, however, it is understood that the method of the invention is suitable for transpeptidation of polypeptides, regardless of the source.

The transpeptidation process of the invention is preferably a one step reaction conducted in a buffer solution capable of maintaining pH at about pH 2-11, preferably pH 3-10, and more preferably pH 5-9. Suitable buffers for the present invention include Mes, Tris, Hepes, Capso, and the like. In one embodiment of the invention using the serine endopeptidase trypsin as the cleavage enzyme for the transpeptidation method of the invention, the modified recombinant polypeptide, for example tagged Glucagon-like Peptide 1 (GLP1) (7-36)-HN $C_6$ $H_3I_2^{128}{}_2$, is produced.

According to the invention, the recombinant polypeptide GLP1 (7-34) core with $Lys^{34}$ linked to the leaving unit -Ala-Phe-Ala with the $Lys^{34}$-$Ala^{35}$ as the cleavage site, is dissolved in buffer. To the transpeptidation mixture is added the suitable addition unit, containing desired amino acid or peptide sequence, in a concentration designed to provide at least a stoichiometric amount, preferably a 100 M excess in an

amount relative to the recombinant polypeptides. For example, Gly-Arg-NH $C_6H_3I_2^{128}$ is desired sequence which is suitable addition units for synthesis of the modified recombinant polypeptide product GLP1 (7-36)-$NHC_6H_3I_2^{128}$. The cleavage enzyme trypsin is added to the reaction mixture in a trypsin:polypeptide molar ration of about 1:2000, and more preferably 1:1000.

The production of the modified recombinant polypeptide product GLP1 (7-36)-$NHC_6H_3I_2^{128}$ is monitored by HPLC or by capillary electrophoresis and the reaction stopped by the addition of an acid solution. Suitable acid solutions for stopping the reaction include HCl and the like.

The modified recombinant polypeptide product is purified from the mixture by HPLC, Ion exchange chromatography, and particle exclusion chain. The recombinant polypeptide product may be used immediately or may be stored at -20 tp -80.

**The Endopeptidase Thrombin Details**

Another example of endopeptidases which may act as a cleavage enzyme according to the method of the invention is thrombin. As described earlier, thrombin has a cleavage site preference for -Arg-X. The known recognition sequence for the -Arg-X cleavage site is Gly-Pro-Arg. A discovery of the present invention is that thrombin also recognizes the cleavage recognition sequence Gly-Ala-Arg. The discovery of this recognition sequence renders the endopeptidase enzyme thrombin of greater utility in preparation of modified proteins by the method of this invention and other recombinant methodologies.

**Thrombin Transpeptidation**

The transpeptidation process of the present invention utilizing the endopeptidase enzyme thrombin is a one pot reaction conducted in a buffer solution

capable of maintaining pH at about pH 2-11, preferably pH 3-10, and more preferably pH 5-9. Suitable buffers for the present invention are as previously described for trypsin. Using the serine endopeptidase thrombin as a cleavage enzyme for cleavage and transpeptidation, the protein includes a GRF (1-41) core linked to an -Ala- leaving unit. A suitable nucleophile for synthesis of GRF (1-44)-Mea is Ala-Arg-Leu-Mea. The present variation uses the discovery that thrombin recognizes the cleavage site -Arg-X within a Gly-Ala-Arg recognition sequence.

This knowledge is used to cleave to the -Ala- leaving unit from the core at the Arg- within Gly-Ala-Arg. The protein GRF (1-41)-Ala is buffer combined with a suitable nucleophile in at least a stoichiometric ratio with the polypeptide. The cleavage enzyme thrombin is added to the mixture in a thrombin:polypeptide molar ratio of about 1:5000, preferably 1:1000, and more preferably 1:3000.

The production of GRF (1-44)-$NH_2$ is monitored by HPLC or other appropriate analytic technique and the reaction stopped by the addition of an acid solution. Suitable acid solutions for stopping the reaction include Hcl, acetic acid and the like. The modified protein is purified from the reaction mixture by standard chromatographic procedure. The protein product may be used immediately or may be stored as a lyophilized product at -80°C.

## Example 1

## Antibody Preparation

Anti-asparagine synthetase monoclonal antibodies and anti-F1 ATPase monoclonal and polyclonal antibodies can be obtained from laboratory stocks. Monoclonal antibody stocks can be obtained in the form of mouse Ascites tumor fluids and polyclonal antibody stocks can be obtained from rabbit serum. Antibodies

from either source can be purified by addition of solid ammonium sulfate to a concentration of 50% of the saturation level. The precipitated protein can be collected by centrifugation and dissolved in a minimal amount of 10 mM Tris-HCl [Tris, tris(hydroxymethyl)aminomethane] (pH 7.5). The preparation can be subjected to a second ammonium sulfate treatment until a 50% saturation level is reached. The precipitate can be collected by centrifugation. The purified antibodies were dissolved in a minimal amount of water and dialyzed for 18 to 24 hours against 10 mM sodium bicarbonate at 4°C. The purified antibodies can be stored in aliquots at -20°C until needed.

The term "antibody" as used herein means anti-asparaginine synthetase monoclonal antibody, and anti-F1 ATPase mono and polyclonal antibodies.

## Example 2

## Label Preparation

The peptide Gly-Asn-$NHC_6H_3I_2^{128}$ can then be prepared by the following technique. The dipeptide Gly-Asn- can be synthesized by chemically protecting the amine group of Gly and coupling to Asn with carbodiimide in DMF. The protected dipeptide can then be amidated with 3,5 diiodo aniline by suitable amidating conditions such as carbodiimide or thionyl chloride.

Radioactive peptides can be diluted with unlabeled amino acid to a specific activity of 0.5 to 2 mCi per millimole. The diluted amino acid is then purified by repeated precipitation with ethanol at -20°C.

The specific activity of the diluted, purified amino acid can be determined as follows, and used in subsequent calculations of label incorporation. A known volume of the amino acid solution can be diluted to 1.0 ml with 0.1 M sodium phosphate (pH 6.8). This diluted

sample can be counted in a Beckman LS-100 liquid scintillation counter (Beckman Instruments, Fullerton, CA) with 10.0 ml 3a70B scintillation fluid (Research Products, Elkgrove, IL). The amino acid concentration of the stock amino acid solution can be determined on a known volume by assay of the amino groups with the ninhydrin assay; see S. Moore et al., J. Biol. Chem., 157, 367 (1948). A standard solution can be prepared for this assay by dissolving glycine in water at 0°C to obtain a saturated solution. The liquid is separated from any undissolved solid glycine, warmed to room temperature, and used as a standard. The concentration of the standard was assumed to be 1.89 M; see J. B. Dalton et al., J. Biol. Chem., 103, 549 (1933). From these measurements, the quenched value of CPM/mmole can be calculated and used in subsequent calculations.

## Example 3

## Standard Conditions for Antibody Labeling

A mixture of a mercury monoclonal antibody made according to the disclosures of U.S. Patent application Serial No. 07/324,392 and Serial No. 493,299, the disclosures of which are incorporated herein by reference and iododipeptide of Example 2 peptide stock solutions can be made and diluted with water to the desired concentrations. From this solution a 4 µl portion can be removed and diluted to 1.0 ml with 0.1 M sodium phosphate (pH 6.8). The concentration of the antibody in the solution can be calculated assuming a standard absorbance of 1.46 absorbance units per mg of antibody; see A. Good et al., in Selected Methods in Cellular Immunology, Mishell & Shiigi eds., W. H. Freeman & Co., San Francisco, p. 284 (1980). The concentration of the peptide in the incubation mixture can be determined by counting a small sample of the diluted mixture following further dilution to 1.0 ml with 0.1 M sodium phosphate (pH 6.8).

The undiluted mixture can be divided into equal parts. To one of these portions can be added the endopeptidase papion and sufficient organic solvent (DMF or 1,4 butanediol to make a 90% solution). The other portion serves as a control and can be diluted similarly with water. Once these two solutions are prepared, they can be incubated in a 37°C water bath for approximately 8 hours. The samples can be then removed from the water bath and either analyzed immediately or frozen and used as soon as possible. The resulting product is $F(ab)_2$ labelled with the iododipeptide.

## Example 4

### Preparation of Amidated Recombinantly Produced GLP1 (7-36)-$NH_2$ From a Single Copy Fusion Protein Construct

A GLP1 peptide is a 36 amino acid peptide that has been recombinantly produced but without a mechanism for providing for the amidation of the C-terminal arginine residue. In this example, the method of the invention has been designed to produce a single copy fusion protein construct containing one copy of a gene encoding a truncated core GLP1 and biotinylating the core GLP1 by a transpeptidation reaction, using the endopeptidase trypsin, to form an immobilizable GLP1 polypeptide.

The strategy involves forming a DNA construct encoding a single copy recombinant fusion protein. The single copy fusion protein includes at least three segments. The first segment is a binding protein which exhibits strong reversible binding to a specific small molecular weight ligand. The second segment is an interconnecting peptide which is selectively cleavable by an enzyme or chemical technique. The third segment is a variable fused peptide containing one copy of the desired natural or synthetic polypeptide, in this case GLP1 (7-34). The formation of a DNA construct for the fusion protein, as well as the fusion protein itself,

has been described in copending U.S. Application Serial No. 07/552,810, which is hereby incorporated by reference.

The single copy fusion protein can be formed with human carbonic anhydrase modified at residue 240 as the binding protein. The modification of carbonic anhydrase at residues 240 involves a substitution of a leucine for a methionine. The interconnecting peptide is a methionine residue which can be cleaved by cyanogen bromide. The variable fused polypeptide contains a single copy of a modified truncated GLP1 (7-37) peptide having Lys at 26, Glu at 27, Lys at 34 and Ala Phe Ala at 35, 36 and 37.

The core GLP1 peptide is truncated from the native sequence so that it contains residues corresponding to residues 7-34 of the naturally occurring sequence. The GLP1 peptide is modified by the linkage of an Ala-Phe-Ala leaving unit at residues 35-37. This tripeptide is not found in the naturally occurring sequence and is a good leaving group for trypsin transpeptidation. Briefly, this single copy recombinant fusion protein can be produced from a DNA construct formed as follows. The DNA sequence from the human carbonic anhydrase II gene is modified so that the methionine codon at amino acid residue 240 is replaced with a leucine codon using site directed mutagenesis as described by Sambrook et al., in Molecular Cloning, "A Laboratory Manual", Cold Spring Harbor, N.Y. (1989). The modified gene for human carbonic anhydrase is then cloned into an expression vector which is compatible with E. coli. A short DNA fragment including the codon for methionine is chemically synthesized and inserted immediately downstream from the end of the gene for human carbonic anhydrase by standard methods. A DNA sequence encoding the truncated core GLP1 (7-34)-Ala-Phe-Ala polypeptide is formed by automated DNA synthesis and inserted directly downstream from the

interconnecting DNA segment encoding the methionine codon. The final recombinant expression vector encoding the single copy fusion protein is transformed into E. coli by standard methods and the expressed recombinant single copy fusion protein can be obtained using affinity chromatography with sulfanilamide or by other chromatographic methods. Once the recombinant fusion protein is purified, it can be cleaved and transpeptidated.

Cleavage and transpeptidation can be conducted as follows. For example, a 40 mg/ml solution of HCA-Met-GLP1 (7-34)-Ala-Phe-Ala can be digested with a 50-fold excess of cyanogen bromide (CNBr) methionine in 70% formic acid to release the GLP1 (7-34)-Ala-Phe-Ala peptide. The reaction mixture can be incubated in the dark under oxygen-free nitrogen at 20°-25°C for 16-24 hours. The reaction mixture is diluted with 15 volumes of water and freeze dried. For the complete removal of acid and by-products, the freeze drying can be repeated after further addition of water. This cleavage reaction yields human carbonic anhydrase and the recombinant GLP1 (7-34)-Ala-Phe-Ala polypeptide.

The cleaved GLP1 7-34 Ala-Phe-Ala polypeptide can be separated from human carbonic anhydrase by normal chromatographic methods, i.e., ion exchange, reverse phase, or by size exclusion. In addition, the cleaved GLP1 (7-34)-Ala-Phe-Ala polypeptide can be separated from the human carbonic anhydrase by simple precipitation procedure. A solution containing carbonic anhydrase, 70% formic acid, cyanogen bromide, methionine, and peptide is diluted with water to a protein concentration of 20 mg/ml while maintaining an acetic acid concentration of 10%. The addition of 5.6 g/100 ml of sodium sulfate ($Na_2SO_4$) to this mixture results in a precipitate which can be removed by centrifugation at 10,000 x g for 10 minutes. The carbonic anhydrase can be quantitatively precipitated

and greater than 80% of the peptide remains in solution. The supernatant can be applied to an open C-8 column which is rinsed with four column volumes of 10% acetic acid. The GLP1 (7-34)-Ala-Phe-Ala can be eluted from the column with 50% acetonitrile in 10% acetic acid. The peptide can then be freeze dried.

Once purified, the GLP1 (7-34)-Ala-Phe-Ala can be transpeptidated to yield the modified recombinant native GLP1 7-36-$NH_2$ amino acid product as follows.

The recombinant GLP1 (7-34)-Ala-Phe-Ala polypeptide can be cleaved with trypsin at the cleavage site between amino acid residues 34 and 35 at the Lys-Als bond in the recombinant truncated polypeptide. Trypsin did not cleave the Lys-Glu bond of residues 26 and 27 in experiments conducted on the recombinant GLP1 polypeptide. While not in any way meant to limit the invention, it is believed that cleavage at residues 26 and 27 by trypsin is not favored because of the presence of the acidic glutamic acid residue.

The cleavage with trypsin is conducted in the presence of a glycyl biotin nucleophiles so that the cleavage of the Ala-Phe-Ala leaving unit is followed by the addition of glycyl biotin.

The biotinylated GLP1 derivative can be bound to an immobilized avidin. The resulting column can be used for an affinity probe for membrane bound GLP receptors.

What is Claimed is:

1. A method for forming a modified protein bound to an auxiliary substance, comprising:

contacting a nucleophile, an endonuclease enzyme specific for an enzyme cleavage site and a protein of at least one leaving unit and a core wherein the leaving unit is linked to the core by the enzyme cleavage site, to form an adduct of the core and the nucleophile, the nucleophile being an amino acid, an amino acid derivative, a peptide or a polypeptide having a side chain with a distinctive reactive substituent; and

binding the adduct to the auxiliary substance or its combination with a linker arm to form the modified protein bound to the auxiliary substance, wherein the auxiliary substance or combination has a specifically reactive group that is correlatively reactive toward the distinctive reactive substituent of the adduct.

2. A method for forming a modified protein bound to an auxiliary substance, comprising:

binding a nucleophile to the auxiliary substance or its combination with a linker arm to form an intermediate, wherein the nucleophile is an amino acid, amino acid derivative, peptide or polypeptide having a side chain with a reactive substituent and the auxiliary substance or combination has a reactive group that is reactive toward the reactive substituent of the nucleophile; and,

contacting together the intermediate, an endonuclease enzyme specific for an enzyme cleavage site and a protein of at least one leaving unit and a core wherein the leaving unit is linked to the

core by the enzyme cleavage site, to form the modified protein bound to the auxiliary substance.

3. A method according to claim 1 or 2 wherein the auxiliary substance is an immobilization support, a label or a bioactive agent.

4. A method according to claim 1 or 2 wherein the combination of auxiliary substance covalently bound to the linker arm having the specifically reactive functional group terminating the free end of the linker arm is used.

5. A method according to claim 1 or 2 wherein the protein is an antibody, an enzyme, an enzyme inhibitor, a protein hormone, a DNA binding protein, a regulatory protein or a DNA reading frame protein.

6. A method according to claim 5 wherein the protein is a polyclonal or monoclonal antibody.

7. A method according to claim 1 wherein the nucleophile is an amino acid selected from the group consisting of an aliphatic amino acid, a hydroxy amino acid, a sulfur-containing amino acid, a diamino monocarboxylic acid, an aromatic amino acid, a heterocyclic amino acid and activated derivatives thereof.

8. A method according to claim 1 or 2 wherein the enzyme is a serine or cysteine endopeptidase.

9. A method for preparing a labeled modified protein, comprising:

contacting together a label nucleophile, and endonuclease enzyme specific for an enzyme cleavage site and a protein of at least one leaving unit and

a core wherein the leaving unit is linked to the core by the enzyme cleavage site, to form the labeled modified protein of the core linked to the label nucleophile.

10. A labeled protein comprising a protein core having transpeptidated to its carboxyl terminus a labeled nucleophilic residue wherein the protein core is derived from a protein starting material of the core and at least one leaving group.

11. A method for detecting the presence of an antigen comprising:

combining a labeled antibody of claim 34 with a material suspected of containing an antigen, to form a complex, said labeled antibody being specifically immunoreactive with said antigen;

removing any uncomplexed labeled antibody; and

measuring the amount of labeled antibody present in the complex.

12. A method according to claim 1 or 2 wherein the auxiliary substance includes a support.

13. An immobilized protein prepared by a method according to claim 1 or 2.

14. A protein augmented by a bioactive agent prepared by a method according to claim 1 or 2.

15. A labeled protein prepared by a method according to claim 1 or 2.

16. A method for functionally reacting a protein comprising employing an immobilized protein of claim 13.

17. A method for conducting a bioreaction with a bioactive agent comprising employing a protein carrying a bioactive agent of claim 14.

18. A method according to claim 1 or 2 wherein the distinctive reactive substituent and specifically reactive group are selected from the pairs consisting of (a) a sulfhydryl group and an organometallic group, (b) an olefinyl group and a dienyl group, (c) a polar olefinic group and its corresponding monomer or substituted forms thereof, (d) an affinity complexing compound and its substrate, (e) a pair of hydroborated olefinic groups, (f) an aromatic amino group and an epoxy activated ester or aldehyde group, (g) an azidyl or hydrazinyl group and an aromatic amino group, (h) an aromatic alcohol and an activated ester group, and (i) a hydrazine and a reducing sugar group.

19. A method according to claim 1 or 2 wherein the linker arm chain is a polymer or oligomer of amide, ester, carbonate, urethane, ether, glycidyl, olefin or hydrocarbon groups, is an aliphatic group of from about 2 to 20 atoms, or is an aromatic group of from about 1 to 5 rings.

20. A method according to claim 1 or 2 wherein the distinctive reactive substituent is selected from the group consisting of sulfhydryl, hydroxyl, activated hydroxyl, olefinyl, activated ester, amino, azidyl, hydrazinyl, phosphoramidoyl, boronyl, ferrocenyl, ferro complexes and mixtures thereof.

21. A method according to claim 1 or 2 wherein the nucleophile is a label and the label is fluorescent, nuclear magnetic, phosphorescent, colorimetric, magnetic, electron resonant or spectrometric.

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 94/08127

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K1/13 C07K1/00 C07K1/107 C12P21/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | EP,A,0 490 249 (HOFFMANN-LA ROCHE) 17 June 1992<br>see the whole document | 1-21 |
| X | EP,A,0 360 433 (R E OFFORD) 28 March 1990<br>see the whole document | 1-21 |
| A | EP,A,0 359 399 (CARLSBERG) 21 March 1990<br>see the whole document | 1-21 |
| A | WO,A,91 00296 (UNIVERSITY OF NEBRASKA/BIONEBRASKA) 10 January 1991<br>see the whole document | 1-21 |

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 9 November 1994 | 24. 11. 94 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax (+31-70) 340-3016 | Masturzo, P |

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/08127

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

see annex

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 94/08127

| Patent document cited in search report | Publication date | Patent family member(s) | | Publication date |
|---|---|---|---|---|
| EP-A-0490249 | 17-06-92 | AU-B- | 644032 | 02-12-93 |
| | | AU-A- | 8896491 | 11-06-92 |
| | | JP-A- | 5178893 | 20-07-93 |
| EP-A-0360433 | 28-03-90 | AU-B- | 637327 | 27-05-93 |
| | | AU-A- | 4204289 | 23-03-90 |
| | | WO-A- | 9002135 | 08-03-90 |
| | | JP-T- | 3501622 | 11-04-91 |
| EP-A-0359399 | 21-03-90 | AU-A- | 4079289 | 05-03-90 |
| | | DE-D- | 68913880 | 21-04-94 |
| | | DE-T- | 68913880 | 22-09-94 |
| | | WO-A- | 9001555 | 22-02-90 |
| | | JP-T- | 4501056 | 27-02-92 |
| WO-A-9100296 | 10-01-91 | US-A- | 5279954 | 18-01-94 |
| | | AU-B- | 646638 | 03-03-94 |
| | | AU-A- | 6030090 | 17-01-91 |
| | | CA-A- | 2063434 | 31-12-90 |
| | | EP-A- | 0479897 | 15-04-92 |
| | | JP-T- | 4506454 | 12-11-92 |
| | | US-A- | 5234820 | 10-08-93 |